

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 04-153878

(43)Date of publication of application : 27.05.1992

---

(51)Int.Cl. G06F 15/38

---

(21)Application number : 02-280336 (71)Applicant : FUJITSU LTD

(22)Date of filing : 18.10.1990 (72)Inventor : IBUKI JUN  
NISHINO FUMITO  
NAKAMURA NAOTO  
SHIOUCHI MASATOSHI  
FUJI HIDE

---

## (54) PRE-EDITING ASSISTING PROCESSOR OF MACHINE TRANSLATION DEVICE

### (57)Abstract:

PURPOSE: To select a proper equivalent by displaying a sentence which is so rewritten that the structure is made clear as to a predetermined sentence pattern in a source sentence as a result of a sentence analysis and passing the rewritten sentence to a machine translating process.

CONSTITUTION: An analytic part 2 for document structure analyzes information regarding the constitution of the whole document which is inputted and sends the result to a sentence analytic part 3 together with the input sentence. The sentence analytic part 3 analyzes the input sentence and transfers part of the analytic result to a structuring process part 4 which converts the received analytic result into a sentence whose structure is clear and displays it at a display part 5.

Over a look at this display a user receives or reject the rewriting result selectively according to the display matches his or her thinking interpretation.

Consequently when the source is vague and possibly translated differently and even when the user does not know the target language the proper equivalent can be selected.

②日本国特許庁(JP)

③特許出願公開

④公開特許公報(A) 平4-153878

⑤Int.Cl.<sup>5</sup>  
G 06 F 15/38

識別記号

序内整理番号

⑥公開 平成4年(1992)5月27日

D  
J 9194-5L  
9194-5L

審査請求 未請求 請求項の数 4 (全7頁)

⑦発明の名称 機械翻訳装置における前編集支援処理装置

⑧特 願 平2-280336

⑨出 願 平2(1990)10月18日

⑩発明者 伊吹潤 神奈川県川崎市中原区上小田中1015番地 富士通株式会社内

⑪発明者 西野文人 神奈川県川崎市中原区上小田中1015番地 富士通株式会社内

⑫発明者 中村直人 神奈川県川崎市中原区上小田中1015番地 富士通株式会社内

⑬発明者 塙内正利 神奈川県川崎市中原区上小田中1015番地 富士通株式会社内

⑭出願人 富士通株式会社 神奈川県川崎市中原区上小田中1015番地

⑮代理人 弁理士森田寛 外2名

最終頁に続く

明細書

機械翻訳処理に引継ぐようにした

ことを特徴とする機械翻訳装置における前編集支援処理装置。

(2) 上記構造化処理部(4)は、原文の中の動詞の連用形と終助詞をもつ文パターンについて、動詞の終止形と接続詞との形に変換するようにしたことを特徴とする請求項①記載の機械翻訳装置における前編集支援処理装置。

(3) 上記構造化処理部(4)は、原文の中の格述の単語列についての評価値を出力する付与体(44)をそなえると共に、当該評価値にもとづいて上記原文を区分するか否かを指示する領収体(46)をそなえたことを特徴とする請求項①記載の機械翻訳装置における前編集支援処理装置。

(4) 上記構造化処理部(4)は、原文における文構造を調べた上で原文を分割し、分割された夫々の文あるいはいずれか一方の文に対して文節を捕らうようにしたことを特徴とする請求項①記載の機械翻訳装置における前編集支援処理装置。

1. 発明の名称

機械翻訳装置における前編集支援処理装置

2. 特許請求の範囲

(1) 与えられた入力文書(1)について文章構造解析部(2)が文章の構造解析を行い、個々の文について文解析部(3)が単語の品詞決定を含む文解析を行って、上記与えられた入力文書(1)における原文を機械翻訳する機械翻訳装置において、

上記文解析部(3)によって行われた文解析の結果にもとづいて、原文の中の予め定められた文パターンについて、より構造が明確にされるよう書き換える構造化処理部(4)と、

当該構造化処理部(4)において行われた結果の書き換え文を表示する表示部(5)とをそなえ、

上記より構造が明確にされた書き換え文を機

特開平4-153878(2)

置。

## 3. 発明の詳細な説明

## (概要)

機械翻訳装置において利用者を支援する機械翻訳装置における前編集支援処理装置に関するもの。

原文があいまいさをもち、翻訳結果に複数の翻訳が得られる可能性がある場合などにおいて、利用者があいまいさのない形に原文を書き換える易くすることを目的とした。

機械翻訳装置における文解析部による解析結果について、予め定められたパターンの文を、より構造が明確にされるような形に書き換える構造化処理部をもつけ、当該構造化処理部による処理結果を表示するよう構成する。

## 〔産業上の利用分野〕

本発明は、機械翻訳装置において利用者を支援する機械翻訳装置における前編集支援処理装置に関するもの。

読みにくいものになることが多かった。この場合、人間の作業者はシステムの出力から使えるような部分訳を書き集め、自分の手で訳を構築するほめになり、作業効率の低下を招いていた。

既に従来、機械翻訳装置から正しい翻訳を得るために、利用者が原文をシステムが正しく解釈できるような形に書き換える必要があった。ところが、その作業はシステムの処理に精通した人が大変な手間をかけて行うものであり、機械翻訳を使った翻訳作業全体のコストを大きく引き上げていた。

この処理を自動的に行うシステムとして自動文分割システムなどが既に発表されているが、実用度の点で問題が多く、実用的なものではなかった。

また訳文の読みやすさに対する配慮は、ほとんどなされておらず、このことと翻訳結果をチェックするのにかかる労力を引き上げ、ひいては翻訳にかかるコストの上昇する要因となっていた。

現在の機械翻訳装置では、意味的情報の欠如、バーザの記述力の不備などのため、文構造の解析を誤りなく行なうことがむずかしい。このため、正しい翻訳結果を得るためには原文を意味さのない、一意に意味の定まるものとする処理（前編集）が必要となる。

## (従来の技術)

現状では、当該前編集の処理に大きな手間がかかり、機械翻訳を利用した翻訳作業の能率の向上の妨げとなっている。これには次の理由が考えられる。

- (1) エディタの操作性が悪く、書き換えに手間取る。
- (2) 翻訳システムの処理内容が明白に規定されていないため、原文をどのように修正すれば翻訳できるのかが判らない。

特に長文などの場合、前編集を行わない文の翻訳成功率は非常に低くなる。また翻訳が成功した場合でも、機械翻訳システムが产出する訳文は非常

## (発明が解決しようとする課題)

初心者がシステムを使う場合、どのような形に翻訳対象文を修正したら正しい翻訳結果が得られるかは、容易にはわからない。そのため、修正なしでシステムにかけ、その結果、翻訳にかなりの誤りが混入するといった問題点があった。一方システムの処理ロジックに精通した人にとっても前編集作業は、文のはとんど全般的な書き換えを必要とするために極めて低能率的なものであった。さらに機械翻訳システムの翻訳結果は必ずしも読みやすいものではなく、翻訳結果を理解することや、翻訳結果を修正して最終的な翻訳を得ることを難しいものとしていた。

本発明は、原文があいまいさをもち、翻訳結果に複数の翻訳が得られる可能性がある場合などにおいて、利用者があいまいさのない形に原文を書き換える易くすることを目的としている。

## (課題を解決するための手段)

第1図は本発明の原理構成図を示す。図中の符

特開平4-153878 (3)

図1は翻訳対象となる入力文書である。2は文章構造解析部であって、入力文書1を例えればタイトル、パラグラフ1、パラグラフ2、…の如く区分するなどの文章構造についての解析を行う。3は文解析部であって、入力文を解析する。4は構造化処理部であって、受けとった解析結果について構造が明確になるような形の文に変換する。5は表示部であって変換結果を表示する。6は文解析部3において得られている別の解析結果を求める別結果探索部を表わす。

本発明では、機械翻訳装置の使用にあたって、原文にあいまいさがある場合に、その可能な解析を原文を書き換えることによって示し、初心者が正しい翻訳を得られるような形に原文を修正する作業をシステムの示した書き換えの可能性の中から選択するだけで行えるように簡単化する。また文の構造がはっきりと示されるように文を書き換えることによって翻訳結果の読みやすさを改善するようとする。

書だけを目指す場合には、対話を用いた別結果探索部5を省略してもよい。

#### (実施例)

第2図は構造化処理部の一実施例構成を示し、第3図はその処理例を示す。第2図において符号4は構造化処理部であり、4-1は構成單語列解析体、4-2は構造化処理体を表わしている。

構成單語列解析体4-1は、入力文を構成する単語を解析し、各単語に対応した品詞を判定する。構造化処理体4-2は予め定められた文パターンを抽出し、当該文パターンをもつ文をより明確な文に変換する。

第3図図示の処理例は、

動詞の連用形 + 終助詞「て」

で与えられる文パターンをもつ文について、

動詞の終止形 + 「。」 + 「そして」

の形に変換する例を表わしている。「処理例1」の場合における「消して」は次行を段落用をとっているものであり、「ファイルを消して処理

#### (作用)

文章構造の解析部2は、入力された文章全体の構成に関する情報を解析して、入力文とともに、文解析部3に送る。文解析部3は、入力文を解析してその解析結果の一部を構造化処理部4に転送する。構造化処理部4は、受けとった解析結果を構造がはっきりと判るような形の文へ変換して表示部5に表示する。各該表示をみてユーザーは自分の考えた解釈と一致するか否かによって書き換え結果を受理するか、拒否するかを選択する。

拒否された場合には、システムは更に別の解析結果を文解析部3から転送し、再び、構造化処理部4の処理にゆだねる。受理する場合には、解析結果はそのまま出力される。この特點ではあいまいな部分についてどのような解釈がとられているかが明らかにされ、ユーザーに承認されていることとなる。

もちろん、上記の場合は、ユーザーとの対話によって翻訳の曖昧さを解決することを目的とした処理の場合である。単に翻訳結果の読みやすさの改

を終える。」という入力文は、「ファイルを消す、そして処理を終える。」という形に変換される。「処理例2」の場合における「消して」は次行実格用法をとっているものであり、「ファイルを消して処理を終える。」という入力文は、「ファイルを消去する、そして処理を終える。」という形に変換される。

この処理の場合、例えば「AしてBしてCする。」の如く連なる場合には、「Aする、そしてBする、そしてCする。」の如く変換されると考えてよい。

第4図は構造化処理部の他の実施例構成を示し、第5図はその処理例を示す。第4図において符号4は構造化処理部であり、4-3は形態素解析体、4-4は特定単語列の検査と評価値の付け体（以下、評価値付与体と略す）、4-5は構造化処理体、5もは構造化処理の制御体を表わしている。

第5図図示の場合には、得られた単語列にもとづいて、評価値付与体4-4が、例えば述語が3個以上存在する如き文を調べる。そして構造化処理

特開平4-153878 (4)

第4は、当該3個以上の場合にあって、述語が各欄以内の文の形に変換するようとする。即ち、評議欄付与体4-4が評議使「述語の個数」を別體4-6に与え、当該別體4-6が「述語3個以上」の条件に合致するか否かをチェックし、構造化処理体4-5に対して、原文を変換するか否かの指示「on/off」の制御信号を与える。

第5図表示の処理例は、「述語が3つ以上存在する」場合に「文の分割化処理を行う」ようにした例を表わしている。

図示例の場合には、「ファイルを消して処理を終える。」という文において、述語が2箇であることから、文の分割化処理は行われない。それに対して、図示例の場合には、「データを退避し、ファイルを消して処理を終える。」という文において、述語が3箇であることから、文の分割化処理が行われる。この例の場合には、「退避し」が各行座標適用であり、「消去し」が各行座標適用であることから、「データを退避し、ファイルを消去する。そして処理を終える。」という形に

第7図(B)における「ユーザは、」の代わりに第7図(C)においては「彼は、」を補なうようにしている。

したがって、第8図に処理例を示す如く、

「ユーザは、ファイルを消去して処理を終える。」

という原文は、

「ユーザは、ファイルを消去する。」

「そして彼は、処理を終える。」

の如く変換されることとなる。第9図表示の場合には第8図表示の場合と同様に変換されるが、併せて単語の品詞をも表示するようにした場合を表わしている。

#### (発明の効果)

以上説明した如く、本発明によれば、原文中にあいまいさがあり、複数の翻訳の可能性がある場合、かつユーザが目標言語を知らない場合においても適切な訳を選択することが可能となる。また本発明によれば、ユーザが機械翻訳システムの基

変換される。

第5図は構造化処理部の更に他の実施例を示し、第7図は文の構造情報を利用した構造化の例、第8図および第9図は夫々処理結果の例を示す。

第6図において、符号4は構造化処理部であり、4-3は形態素解析体、4-7は文構造解析体、4-8は構造化処理体本体を表わしている。

第6図表示の場合には、第7図を参照して説明する如く、文構造解析体4-7が文中の語彙関係を明らかにし構造化処理体本体4-8が原文を変換するようしている。

第7図(A)は原文の「ユーザは、ファイルを消去して処理を終える。」について文構造を明らかにした状態を表わしている。図中のSは文、Pは述語、P-Pは文節を表わしている。第7図(A)においては、(A)「ファイルを消去して」の部分と、(B)「処理を終える。」の部分とが、同じ「ユーザは、」の述語に対応している。この結果、第7図(B)と第7図(C)との如く、2つの文に分割し、第7図(C)の場合の如く、第

理解できる範囲の文についての知識がなく、目標言語を知らない場合においても適切な訳を選択することが可能となる。

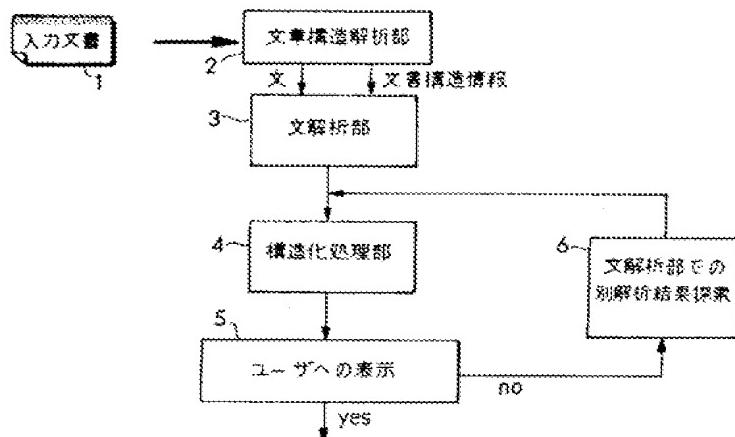
なお本発明においては、上記第2図、第4図、第6図表示の各構成のいずれか1つのみの場合であってよいことは言うまでもなく、いずれか複数個の組合せをもつ場合であってもよい。

#### 4. 図面の簡単な説明

第1図は本発明の原理構成図、第2図は構造化処理部の一実施例、第3図は第2図表示の場合の処理例、第4図は構造化処理部の他の実施例、第5図は第4図表示の場合の処理例、第6図は構造化処理部の更に他の実施例、第7図は文の構造情報を利用した構造化の例、第8図および第9図は夫々処理結果の例を示す。

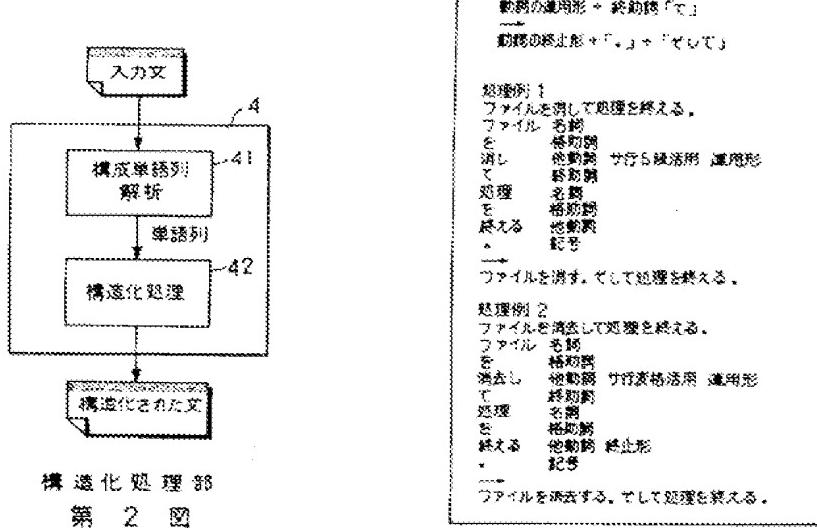
図中、1は入力文書、2は文章構造解析部、3は文解析部、4は構造化処理部、5は表示部、6は別結果探索部を表わしている。

特開平4-153878(5)



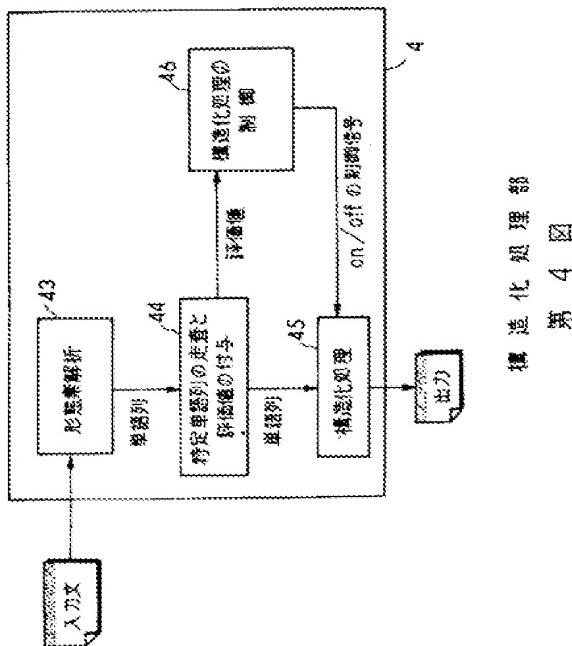
本発明の原理構成図

第 1 図

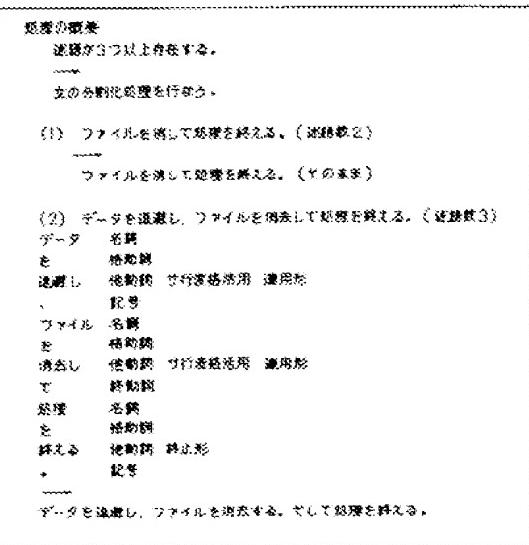
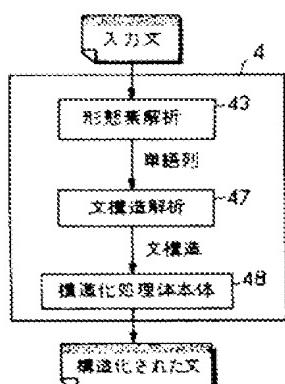
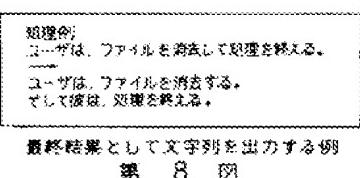
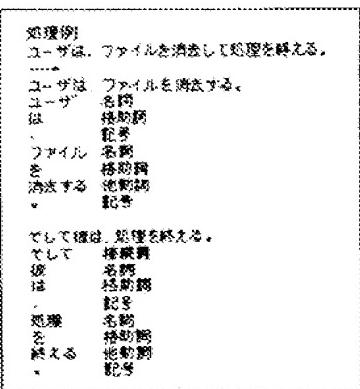
構造化処理部  
第 2 図

第 3 図

特開平4-153878 (6)

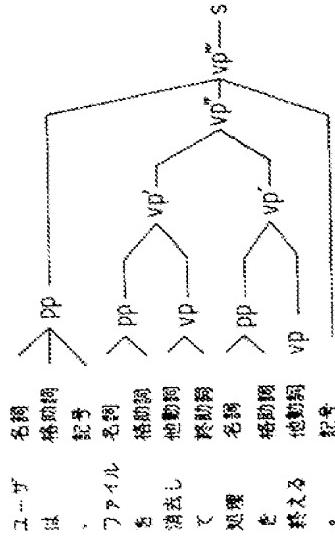


構造化処理 第4図

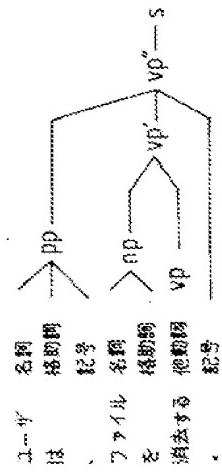
処理例  
第5図構造化処理部  
第6図最終結果として文字列を出力する例  
第8図最終結果として単語列を出力する例  
第9図

特開平4-153878(7)

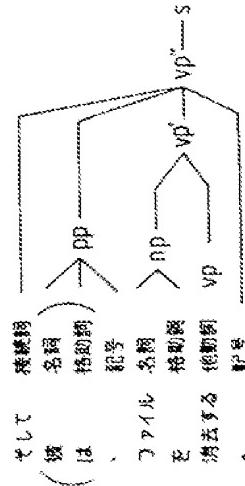
(A) ユーザは、ファイルを消去して処理を終える。



(B) ユーザは、ファイルを消去する。



(C) そして彼は処理を終える。

文構造の例  
第7図

第1頁の続き

◎発明者　齋　洋　秀　神奈川県川崎市中原区上小田中1015番地　富士通株式会社